MATH 321 – SPRING 2019 – EXAM 3 SOLUTIONS

1. (12 pts) A steel cable manufacturer wants to estimate the mean strength of their product for quality control purposes. A sample of 7 sections of cable are pull-tested until failure. The mean force required to cause failure in the sample is found to be 12,500 lbs with a sample standard deviation of 800 lbs. Calculate a 99% lower bound on true mean force required to cause cable failure. Writing the appropriate R code is acceptable. The numbers are not nice here.

   Solution:

   $12500 - \texttt{qt(0.99,df=6)}\frac{800}{\sqrt{7}}$

2. (12 pts) An electronics component manufacturer wishes to estimate the mean lifespan of its resistors under a high voltage load. A sample of 400 resistors is tested and the mean time until failure is found to be 137 minutes with standard deviation 40 minutes. Estimate the mean time to failure with a 95% confidence interval. Do not use R code, estimate the interval by hand.

   Solution:

   The sample is very large, so we can use the CLT normal approximation. Also, ± 1 standard deviations gives a z quantile of about 2 by the 68-95-99.7 rule.

   $137 \pm \texttt{qnorm(0.975)}\frac{40}{\sqrt{7400}} \approx 137 \pm 2 \cdot \frac{40}{20} = 137 \pm 4 = (133, 141)$

3. (12 pts) A political polling company surveys two regions regarding support for a particular ballot proposition. Out of 300 respondents in region A, 210 of them will vote 'yes.' Out of 300 respondents in region B, 150 state they will vote 'yes'. Conduct a statistical hypothesis test at significance level $\alpha = 0.05$ about whether or not there truly is a difference in the support for the ballot proposition between the two regions. **Calculate and simplify everything by hand. The numbers are nice if you use the pooled proportion.**

   Solution:

   H_0 : the proportions are equivalent, $p_1 = p_2$
   $H_a : p_1 \neq p_2$ (a 1-sided test would also be acceptable)

   The pooled proportion is $\hat{p} = \frac{210+150}{300+300} = 0.6$

   The test statistic is

   $$z = \frac{0.7 - 0.5}{\sqrt{\frac{6}{10}\frac{4}{10}\frac{2}{300}}} = \frac{0.2}{\frac{4 \cdot 4}{100^2}} = \frac{0.2}{4/100} = 0.2 \cdot \frac{100}{4} = 5$$

   This is a very large z value, well beyond the 2 standard deviations associated with the 5% level for a 2-sided test. Thus we reject H_0. This conclusion holds regardless of 1- vs 2- sided.

4. A scientist at a water treatment plant takes 9 samples to test for a particular contaminant. The current regulatory agency allowed level is 5 parts per million (ppm) for this particular contaminant. The mean contaminant level is found to be $\overline{X} = 5.4$ ppm with a sample standard deviation of 0.1 ppm. A statistical hypothesis test is to be conducted with significance level $\alpha = 0.05$.

   (a) (6 pts) Write down null and alternative hypotheses.

   Solution: $H_0 : \mu = 5$, $H_a : \mu > 5$

(b) (6 pts) Calculate the needed test statistic. Simplify exactly by hand.

Solution: $t = (5.4 - 5)/(0.1/\sqrt{9}) = 12$

(c) (6 pts) Calculate the p-value. You may write R code here.

Solution: $p =$ `1-pt(12,8)`

(d) (6 pts) State your conclusion. Is there strong evidence that the water is or is not within regulatory guidelines? Even if you do not know the p-value exactly, use your knowledge to draw a definite conclusion.

Solution:

Our test statistic t is very large, that it will give a very small p-value, thus we will likely reject the null hypothesis with strong evidence.

5. (9 pts) Calculate the p-value for the following hypothesis tests for the mean of a normally distributed population and unknown variance. Answering in terms of R code is acceptable.

(a) $H_0 : \mu = \mu_0$
$H_a : \mu \neq \mu_0$
and test statistic $t = 1.76$ and sample size $n = 20$.

Solution: This is 2-sided, so either `2*pt(-1.76,19)` or `2*(1-pt(1.76,19))`

(b) $H_0 : \mu = \mu_0$
$H_a : \mu > \mu_0$
and test statistic $t = 1.87$ and sample size $n = 45$.

Solution:

either `pt(-1.87,44)` or `1-pt(1.87,44)` are fine

or since we have a large sample size, we can use normal:

`pnorm(-1.87)` or `1-pnorm(1.87)`

(c) $H_0 : \mu = \mu_0$
$H_a : \mu < \mu_0$
and test statistic $t = -2.15$ and sample size $n = 12$.

Solution:

either `pt(-2.15,11)` or `1-pt(2.15,11)` are fine

6. (9 pts) Calculate the p-value for the following hypothesis tests for the mean of a normally distributed population and known variance. Answering in terms of R code is acceptable.

(a) $H_0 : \mu = \mu_0$
$H_a : \mu \neq \mu_0$
and test statistic $z = 2.76$.

Solution: `2*pnorm(-2.76)` or `2*(1-pnorm(2.76))`

(b) $H_0 : \mu = \mu_0$
$H_a : \mu > \mu_0$
and test statistic $z = 1.79$.

Solution: `pnorm(-1.79)` or `1-pnorm(1.79)`

(c) $H_0 : \mu = \mu_0$

$H_a : \mu < \mu_0$

and test statistic $z = -1.15$.

<u>*Solution:*</u> `pnorm(-1.15)` or `1-pnorm(1.15)`

7. (10 pts) Consider the dataset below. Calculate by hand: the least squares regression line, the correlation coefficient, and sum of squared errors. Hint: the numbers are nice!

x	10	12	20	30	36
y	2050	2060	2100	2150	2180

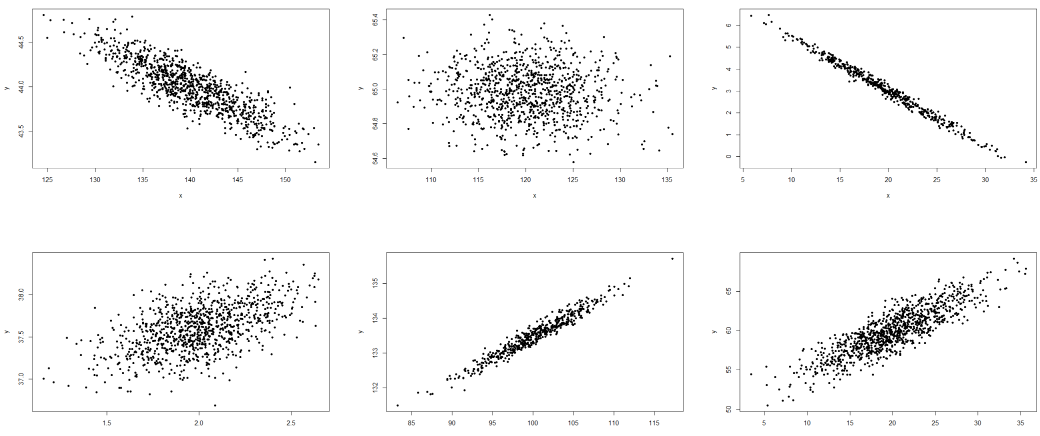
<u>*Solution:*</u>

Data is exactly linear. $y = 2000 + 5x$ so $R = 1$ and $SSE = 0$.

The other exam version had: Data is exactly linear. $y = 1000 + 2x$ so $R = 1$ and $SSE = 0$.

8. (12 pts) Match the correlation coefficient to the graph.

$R = -0.99$, $R = -0.06$, $R = 0.86$, $R = 0.54$, $R = 0.97$, $R = -0.85$,



<u>*Solution:*</u>

In order:

$R = -0.85, -0.063, -0.99$

$R = 0.54, 0.97, 0.86$